



Approches simultanées et séquentielles de la mise en correspondance par propagation

Guillaume Gales, Alain Crouzil, Sylvie Chambon

► To cite this version:

Guillaume Gales, Alain Crouzil, Sylvie Chambon. Approches simultanées et séquentielles de la mise en correspondance par propagation. 12ème Congrès des jeunes chercheurs en vision par ordinateur (ORASIS 2009), IRISA : Institut de Recherche en Informatique et système Aléatoires; INRIA Rennes-Bretagne Atlantique; AFRIF: Association Française pour la Reconnaissance et l'Interprétation des Formes, 2009, Trégastel, France. inria-00404642

HAL Id: inria-00404642

<https://inria.hal.science/inria-00404642>

Submitted on 16 Jul 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Approches simultanées et séquentielles de la mise en correspondance par propagation

Simultaneous and sequential approaches for propagation-based matching

Guillaume Gales¹

Alain Crouzil¹

Sylvie Chambon²

¹ Institut de Recherche en Informatique de Toulouse (IRIT)

² Laboratoire Central des Ponts et Chaussées (LCPC)

¹ 118 route de Narbonne, 31062 Toulouse Cedex 9

² route de Pornic, BP 4129, 44341 Bouguenais Cedex
gales, crouzil@irit.fr ; sylvie.chambon@lcpc.fr

Résumé

La mise en correspondance de pixels est une étape importante de la reconstruction 3D. Parmi les méthodes existantes, nous nous intéressons plus particulièrement à celles par propagation de germes qui s'appuient sur un ensemble d'appariements fiables (germes). Le principe utilisé pour la propagation est que le correspondant d'un voisin d'un germe est recherché au voisinage du correspondant de ce germe. Il existe deux types d'approches pour ces méthodes : l'approche séquentielle et l'approche simultanée. Dans cet article, nous proposons un état de l'art de ces méthodes ainsi qu'une amélioration de l'approche simultanée. Notre méthode se déroule en deux phases faisant intervenir deux mesures de corrélation dont une est robuste aux occultations dans les zones proches des discontinuités de profondeur. Nous proposons également une étude comparative des différentes approches.

Mots Clef

Mise en correspondance de pixels, propagation, germes.

Abstract

Pixel matching is an important step in 3D reconstruction. Among all existing methods, we take a closer look at the ones using seed propagation. They start from a set of reliable matches (seeds) and look for a correspondent of the neighbors of a seed in the neighborhood of this seed correspondent. There are two kinds of approaches: sequential and simultaneous. In this paper, we give a state of the art of these methods and we propose to improve the simultaneous one. This is done in two steps using two correlation measures. One of these is robust to occlusions near depth discontinuities. We also propose a comparative study of these different approaches.

Keywords

Pixel matching, propagation, seeds.

1 Introduction

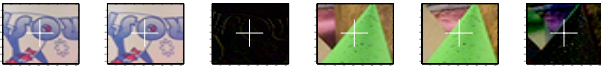
La mise en correspondance de pixels est une étape importante de la reconstruction 3D dont les applications sont très nombreuses (robotiques, industrielles). Le problème consiste à trouver pour chacun des pixels de l'image de référence (ou d'une partie de l'image de référence), le pixel qui correspond à la même entité de la scène sur l'autre image. Il existe deux grandes classes de méthodes qui cherchent à résoudre ce problème : les méthodes globales et les méthodes locales. Des états de l'art de référence existent dans [1, 15]. Parmi les méthodes locales, nous nous intéressons aux méthodes par propagation de germes.

Les méthodes globales prennent en compte la totalité des pixels étudiés pour estimer la validité de l'ensemble des correspondances. L'idée générale est la suivante : si on « déplace » les pixels de l'image de gauche suivant des valeurs de disparités, alors si ces valeurs sont correctes, l'image obtenue doit être identique à l'image de droite. Par conséquent, on cherche à minimiser une fonction de coût global (coût sur l'ensemble de toutes les correspondances) qui intègre un terme d'attache aux données représentant la différence entre l'image de gauche « transformée » et l'image de droite, ainsi qu'un terme de régularisation. Des techniques d'optimisation (coupure de graphe, recuit simulé, propagation de croyance) sont habituellement utilisées pour trouver un tel minimum et donnent de bons résultats mais sont généralement coûteuses en temps de calcul. À titre d'exemple, nous pouvons citer la méthode de [8] qui utilise la propagation de croyance et qui est classée première au classement de Middlebury¹ dont le protocole d'évaluation est présenté dans [15].

Les méthodes locales s'appuient sur des mesures de corrélation. Ces mesures évaluent localement la similarité de

¹vision.middlebury.edu/stereo/Eval/

deux ensembles de pixels. Pour un pixel et son voisinage dans l'image de référence, on cherche dans l'autre image le pixel et son voisinage qui leur ressemblent le plus. Un état de l'art sur les mesures de corrélation se trouve dans [2]. Ces méthodes sont simples à mettre en œuvre et moins coûteuses que les méthodes globales mais peuvent être confrontées à certaines ambiguïtés. De plus, un problème se pose aux niveaux des occultations où un pixel de l'image de référence n'a pas de correspondant visible dans l'autre image, et/ou des discontinuités de profondeur où les voisinages d'un pixel et de son correspondant risquent d'être différents, voir figure 1. C'est pourquoi des mesures robustes ont été proposées dans [2] où seulement la partie la plus fiable du voisinage est prise en compte.



(a) gauche (b) droite (c) diff. (d) gauche (e) droite (f) diff.

FIG. 1 — Ces sous-ensembles appartiennent au couple *cones* (voir 4.2). Sur cet exemple, (a) montre le voisinage du pixel marqué d'une croix blanche et (b) le voisinage de son correspondant. Ces deux ensembles n'incluent pas de zone de discontinuités. Les niveaux de gris sont quasi-identiques. L'image (c) montre qu'il n'y a pas (ou très peu) de différences entre eux. Une mesure de corrélation entre (a) et (b) donne un bon score. L'image (d) montre le voisinage d'un autre pixel et (e) le voisinage de son correspondant. Ces deux ensembles incluent des zones de profondeurs différentes. Les niveaux de gris de l'arrière-plan étant différents, une mesure de corrélation classique ne donne pas un bon score. L'image (f) montre la différence entre ces deux ensembles. Une mesure de corrélation robuste est alors préconisée pour limiter les méfaits d'une telle configuration.

Dans [17, 18, 20], les auteurs proposent des méthodes progressives. Elles se déroulent en plusieurs phases. Par exemple, l'idée de [17, 20] est de réaliser d'abord les appariements les plus fiables, réduisant ainsi les zones de recherche des pixels pour lesquels une valeur de fiabilité n'est pas assez élevée et qui pourront être appariés lors des itérations suivantes. Dans [18], les auteurs proposent d'effectuer les appariements les plus fiables dans une première phase, puis de segmenter l'image en régions et d'attribuer des disparités à ces régions en fonction des appariements déjà obtenus.

Les méthodes par propagation de germes s'appuient sur un ensemble d'appariements fiables, les germes, et se fondent sur l'hypothèse que le correspondant d'un voisin d'un germe se situe au voisinage du correspondant du germe. Les correspondances trouvées au voisinage des germes servent de germes à leur tour et le processus est répété jusqu'à l'obtention d'une carte de disparités dense ou quasi-dense. La zone de recherche des correspondants est ainsi minimale, ce qui augmente la vitesse d'exécution ainsi que la fiabilité des résultats. Nous pouvons distinguer deux approches pour la propagation : une approche séquentielle, comme dans [7, 9, 11], où les germes propagés sont sélectionnés dans un ordre spécifique, et une approche simulta-

née, comme dans [3, 20], qui va produire un « effet tâche d'huile » (la propagation a lieu dans le voisinage des correspondants fiables). Nous proposons une amélioration de cette méthode en distinguant deux phases de propagation faisant intervenir deux mesures de corrélation dont une est robuste dans les zones proches des discontinuités de profondeur.

Dans un premier temps nous donnons un état de l'art des différentes méthodes par propagation de germes puis nous décrivons l'approche simultanée en deux phases. Nous proposons également une évaluation et une comparaison des différentes méthodes avant de conclure.

2 Méthodes par propagation

Nous nous intéressons dans cette section aux différentes méthodes de propagation de germes. De telles méthodes peuvent se découper selon les étapes suivantes :

1. sélection des germes ;
2. mise en correspondance des germes ;
3. mise en correspondance au voisinage des germes (propagation).

Les appariements trouvés lors de ce dernier point peuvent alors être considérés comme de nouveaux germes et cette étape est réitérée jusqu'à l'obtention d'une carte de disparités dense ou quasi-dense. Ces nouveaux germes peuvent être utilisés de façon simultanée (approche simultanée) comme dans [3, 20] ou bien un seul germe peut-être sélectionné de façon spécifique (approche séquentielle), par exemple en choisissant celui qui a obtenu le meilleur score de corrélation comme dans [7, 9].

2.1 Définitions et notations

On note $\mathbf{p}_{i,j}^k$ le pixel situé à la ligne i et à la colonne j de l'image k avec $k \in \{g, d\}$ où g fait référence à l'image gauche et d à l'image droite. Le correspondant d'un pixel est donné par la fonction de disparité \mathcal{D} qui représente son « déplacement apparent ». Nous nous plaçons dans le cadre d'images stéréoscopiques à base étroite, c'est-à-dire prises de point de vue qui ne sont pas trop éloignés. Le terme $\mathbf{f}_{i,j}^k$ représente le vecteur contenant l'ensemble des niveaux de gris au voisinage du pixel $\mathbf{p}_{i,j}^k$. Cet ensemble est appelé voisinage ou fenêtre de corrélation. Pour plus de lisibilité, on appelle s le germe considéré à un instant t dans l'image de référence et s' son correspondant dans l'autre image.

2.2 Sélection des germes

Il s'agit dans un premier temps de sélectionner un ensemble d'appariements fiables. Nous pouvons distinguer les méthodes qui utilisent des correspondances de points quelconques de l'image et celles qui calculent des points d'intérêt (pixels qui possèdent des caractéristiques permettant de les distinguer des autres pixels de l'image).

Dans [18], les auteurs considèrent tous les pixels de l'image comme des germes potentiels. Seuls les appariements les plus fiables sont conservés. La cardinalité de

l'ensemble des germes est alors maximale au détriment d'un temps de calcul plus long.

Dans [3], les auteurs proposent une technique de baquets. L'image est divisée en blocs parmi lesquels est tiré au sort un pixel pour lequel on cherche un correspondant. Si ce correspondant ne peut pas être déterminé de façon fiable, alors un autre pixel est tiré au sort. Le processus est répété tant que tous les blocs ne disposent pas d'un appariement fiable et que tous les pixels de chaque bloc n'ont pas été testés sans succès. La répartition des germes est relativement uniforme mais le tirage au sort ne garantit pas que les appariements trouvés soient les meilleurs en terme de fiabilité.

Dans [20], les auteurs proposent de faire une sélection manuelle des germes. Le contrôle de l'opérateur est alors absolu mais pas exempt d'erreurs.

Les points d'intérêt du détecteur de Harris [5] sont utilisés dans [9, 17, 20]. Le détecteur Hessienne-Affine de [12] est utilisé dans [7]. Ces points sont *a priori* plus faciles à mettre en correspondance que les autres sauf lorsqu'ils sont situés au niveau des discontinuités de profondeur.

2.3 Mise en correspondance des germes

La mise en correspondance des germes peut-être réalisée en utilisant les techniques classiques de mise en correspondance par corrélation, ou bien, dans le cas de points d'intérêt, en utilisant des descripteurs spécifiques.

Les mesures de corrélation sont généralement utilisées pour la recherche d'un correspondant d'un point donné. Pour chaque point de l'image de référence dont on cherche le correspondant, on calcule les scores de corrélation avec les points de l'autre image qui constituent la zone de recherche. Cette zone de recherche est dans ce cas relativement étendue. Il peut s'agir de l'image dans son intégralité si nous n'avons aucune connaissance *a priori*. Elle peut-être réduite à une seule ligne, en particulier si les images ont subi une rectification épipolaire. Le candidat de la zone de recherche ayant donné le meilleur score est considéré comme correspondant potentiel. Lors de l'étape suivante de sélection, ce correspondant peut être rejeté s'il ne respecte pas certaines contraintes.

Les mesures de corrélation classiques telles que la somme des valeurs absolues des différences des niveaux de gris SAD (*Sum of Absolute Differences*) ou la somme des carrés des différences des niveaux de gris SSD (*Sum of Squared Differences*) sont utilisées dans [17]. Des mesures normalisées comme NCC (*Normalized Cross Correlation*) ou ZNCC (*Zero-mean Normalized Cross Correlation*) sont utilisées dans [3, 9, 11, 20]. La mesure de corrélation croisée centrée normalisée ZNCC est définie par :

$$\text{ZNCC}(\mathbf{f}_{i,j}^g, \mathbf{f}_{i',j'}^d) = \frac{(\mathbf{f}_{i,j}^g - \overline{\mathbf{f}_{i,j}^g})^T (\mathbf{f}_{i',j'}^d - \overline{\mathbf{f}_{i',j'}^d})}{\|\mathbf{f}_{i,j}^g - \overline{\mathbf{f}_{i,j}^g}\| \|\mathbf{f}_{i',j'}^d - \overline{\mathbf{f}_{i',j'}^d}\|}$$

où $\overline{\mathbf{f}_{i,j}^g}$ est un vecteur contenant la moyenne des niveaux de gris des $\mathbf{f}_{i,j}^k$ et $\|\cdot\|$ est la norme euclidienne. Cette mesure

a l'avantage d'être invariante aux changements linéaires de luminosité entre les deux images. Les valeurs de ZNCC sont comprises entre $[-1; +1]$. Une fenêtre de corrélation trop petite augmente la sensibilité au bruit mais une fenêtre trop grande risque de couvrir des zones de profondeurs différentes et donc de ne pas donner un meilleur score pour l'appariement juste. Des fenêtres de taille variable peuvent alors être utilisées pour pallier ce problème [11].

Dans le cas d'une mise en correspondance de germes par corrélation, des contraintes supplémentaires sont souvent utilisées afin d'éliminer les appariements les moins fiables. La contrainte que nous trouvons le plus souvent dans la littérature consiste à fixer un seuil sur le score de corrélation. Elle est utilisée afin d'éliminer les correspondances entre les ensembles très peu similaires, comme dans [3, 9, 18]. Toutefois, un seuil sur le score de corrélation ne suffit pas toujours à assurer la fiabilité d'un appariement en particulier dans les zones homogènes où les pixels ont des niveaux de gris très proches. Une valeur d'ambiguïté peut alors être calculée comme dans [3, 17]. Elle mesure pour un point donné si son correspondant potentiel donne un score qui se distingue fortement des scores des autres candidats. Par exemple, dans [18], le rapport du meilleur score sur le second meilleur score est utilisé. Cette étape est utile lorsque l'on recherche des correspondants fiables mais risque d'éliminer des appariements corrects. La vérification bidirectionnelle est aussi largement employée, comme dans [3, 9, 11]. Si on trouve $\mathbf{p}_{i',j'}^d$ comme correspondant de $\mathbf{p}_{i,j}^g$, alors on doit retrouver ce même $\mathbf{p}_{i,j}^g$ lorsque l'on recherche le correspondant de $\mathbf{p}_{i',j'}^d$.

Lorsque les points d'intérêt sont utilisés, des descripteurs spécifiques peuvent remplacer les niveaux de gris des voisinages pour caractériser un point d'intérêt et son voisinage. Ils sont généralement utilisés dans des applications d'indexation. Certains sont robustes à différentes transformations telles que la rotation ou le changement d'échelle [10, 12]. Le descripteur SIFT (*Scale Invariant Features Transform*) de [10] est utilisé dans [7]. Il est calculé à partir des histogrammes des orientations du gradient au voisinage du point d'intérêt considéré.

2.4 Propagation

Une fois l'ensemble des germes initiaux mis en place, on passe à l'étape de propagation de ces germes. On distingue deux approches. L'approche simultanée et l'approche séquentielle. La première va réaliser une propagation de tous les germes présents à une même itération en même temps. La seconde, l'approche séquentielle, ne va propager qu'à partir d'un seul germe qui aura été sélectionné suivant certains critères. En fonction des différents cas d'arrêt et des différentes contraintes, la propagation peut être dense ou quasi-dense.

Approche simultanée. Dans [3, 19], lors d'une même itération, tous les germes présents à l'instant t vont avoir la même importance et peuvent être propagés simultanément. Cela va produire un « effet tâche d'huile ». Cette approche

a toutefois l'inconvénient d'être sensible aux erreurs de départ. Dans la version originale de l'article [3], un algorithme *a priori* séquentiel est proposé, mais il peut se ramener à une approche simultanée car les auteurs remettent en question les appariements déjà effectués si un appariement meilleur est trouvé lors d'une itération future. L'algorithme peut se résumer suivant le modèle général présenté dans l'algorithme 2.1.

```

Entrée :  $\mathcal{G}$ , ensemble des germes initiaux
Sortie :  $\mathcal{A}$ , ensemble des appariements
Tant que ! cas d'arrêt
1.  $\mathcal{G}_{\text{local}} \leftarrow \emptyset$ 
2. Pour chaque  $s$  de  $\mathcal{G}$ 
  2.1. supprimer  $s$  de  $\mathcal{G}$ 
  2.2. Pour chaque  $p$  du voisinage de  $s$ 
    2.2.1. rechercher son correspondant au
           voisinage de  $s'$ 
    2.2.2. Si cet appariement  $a$  vérifie un
           ensemble de conditions
      A. Si ce pixel a déjà un appariement
          $a_2$  et si  $a$  est meilleur que  $a_2$ ,
         supprimer  $a_2$  de  $\mathcal{A}$ 
      B. ajouter  $a$  dans  $\mathcal{A}$  et  $\mathcal{G}_{\text{local}}$ 
3.  $\mathcal{G} \leftarrow \mathcal{G}_{\text{local}}$ 

```

ALGO. 2.1 – Algorithme générique de l'approche simultanée.

Approche séquentielle. Dans [7, 9, 11], l'ordre dans lequel les germes vont être employés a une influence sur le résultat. Dans [11], cette influence est implicite. En effet, lorsqu'un appariement est trouvé, il n'est pas remis en cause, par conséquent, selon l'ordre de parcours des germes, on risque de ne pas trouver le même résultat. Dans [7, 9], lors de chaque itération, on sélectionne le germe ayant le meilleur score de corrélation. Les auteurs parlent de « stratégie meilleur d'abord ». Les pixels ayant déjà un correspondant ne sont plus remis en cause. Cette approche est moins sensible aux erreurs de départ. L'algorithme peut se résumer suivant le modèle général présenté dans l'algorithme 2.2.

```

Entrée :  $\mathcal{G}$ , ensemble des germes initiaux
Sortie :  $\mathcal{A}$ , ensemble des appariements
Tant que ! cas d'arrêt
1. sélectionner un germe  $s$  de  $\mathcal{G}$  suivant un critère
   prédéfini
  1.1. supprimer  $s$  de  $\mathcal{G}$ 
  1.2. Pour  $p \in$  voisinage de  $s$ 
    1.2.1. rechercher son correspondant potentiel
           au voisinage de  $s'$ 
    1.2.2. Si cet appariement  $a$  vérifie un
           ensemble de conditions, ajouter  $a$  dans
            $\mathcal{A}$  et  $\mathcal{G}$ 

```

ALGO. 2.2 – Algorithme générique de l'approche séquentielle.

La méthode employée pour la recherche des correspondants des voisins d'un germe n'est pas spécifique à une des deux approches. Il s'agit des mêmes méthodes de corrélation classiques décrites précédemment pour la mise en correspondance des germes. Cependant, les paramètres et les contraintes peuvent être différents. C'est le cas de la zone

de recherche réduite au voisinage direct (8-voisins ou 4-voisins) du correspondant du germe considéré. Un tel voisinage garantit alors que le gradient de disparité est limité à 1. Dans le cas d'images ayant subi une rectification épipolaire, ce voisinage peut encore être réduit au 2-voisins de la même ligne que le correspondant du germe, c'est le cas dans [3, 11, 20] et dans [9] pour la propagation contrainte. Dans [9], l'auteur propose également une propagation non-contrainte qui donne de bons résultats sur des images non-rectifiées. L'ensemble des appariements alors trouvés est utilisé pour l'estimation de la matrice fondamentale de manière robuste. Une contrainte sur les seuils des scores de corrélation est employée dans [3, 7, 9, 20]. Dans [7, 9] une contrainte d'homogénéité est utilisée : les pixels dont les voisins ont des niveaux de gris trop homogènes ne sont pas propagés. Cela a pour effet de réduire le nombre d'erreurs mais réduit également la densité.

Dans [20], les auteurs proposent une vérification bidirectionnelle moins stricte que la méthode classique. En effet, si la même valeur de disparité n'est pas retrouvée selon le sens de mise en correspondance, et si la différence entre les deux n'est pas plus grande qu'un seuil, ils proposent d'attribuer au pixel une valeur de disparité qui correspond à la moyenne des deux disparités pondérées par les scores de corrélation obtenus. Dans [9], une vérification bidirectionnelle n'est pas réalisée de façon implicite mais l'unicité des appariements est garantie en ne prenant pas en compte les pixels déjà appariés lors de la propagation.

Dans [11], une contrainte de région est utilisée. Elle impose que pour un germe, seuls les voisins appartenant à la même région que ce germe soient considérés pour la propagation limitant ainsi les erreurs de propagation.

3 Approche multi-mesure

3.1 Objectifs

Nous proposons d'étudier une approche de propagation qui se déroule en plusieurs phases faisant intervenir plusieurs mesures de corrélation adaptées à la situation, voir figure 2. Dans notre étude, deux mesures de corrélation interviennent. La première phase fait appel à une mesure de corrélation classique et recherche les appariements dans les parties de la scène qui posent les moins de problèmes (principalement dans les zones bien texturées et sans occultations et/ou sans discontinuités de profondeur). La seconde phase poursuit la propagation à l'aide d'une mesure robuste. Dans un premier temps, l'approche simultanée nous semble être la plus naturelle pour tester notre proposition de par son effet « tâche d'huile » qui va propager des germes vers les zones difficiles à appairer.

3.2 Initialisation

Segmentation. Une segmentation est calculée en initialisation. Elle est utilisée pour la contrainte de région lors de la propagation, voir 3.3. On suppose que tous les pixels d'une même région appartiennent à la même facette d'un objet et donc ne présentent pas de discontinuités de profon-

deur. Intuitivement, l'utilisation d'une contrainte de région doit donc permettre d'éviter de propager dans des régions connexes de profondeur différente et limiter les erreurs de propagation. Ces régions sont obtenues à l'aide de la méthode de segmentation couleur *mean-shift* de [4].

Calcul des germes initiaux. La sélection et la mise en correspondance des germes a une influence sur le résultat mais dans la mesure où le centre d'intérêt de ce papier porte sur les techniques de propagation, nous nous appuyons dans un premier temps sur un ensemble théorique de germes commun à toutes les méthodes utilisées pour les tests. La réponse du détecteur de Harris [5] est utilisée pour obtenir un ensemble de points d'intérêt. Ces pixels sont ensuite mis en correspondance en utilisant les cartes de disparités théoriques fournies avec les couples stéréoscopiques que nous utilisons pour nos tests (voir section 4.2).

3.3 Propagation

Recherche des correspondants candidats. La première passe est réalisée en utilisant une mesure de corrélation classique, avec ZNCC. Une contrainte de seuil ne conserve que les meilleurs appariements, voir figure 2. On constate alors que la propagation s'arrête avant d'atteindre les zones de discontinuités de profondeur, voir figure 3. Cela s'explique par le fait que les scores de corrélation obtenus lorsqu'on est proche d'une telle zone ne respectent pas la contrainte de seuil et ne sont pas validés. En effet, à ce niveau là, les voisinages entre un pixel sur l'image de gauche et son correspondant sur l'image de droite sont différents, voir figure 1. C'est pourquoi nous proposons une seconde phase faisant intervenir une mesure robuste. Le principe repose sur le fait que les pixels qui ont une profondeur différente du pixel étudié sont considérés comme des données aberrantes. Nous nous basons alors sur l'estimateur SMAD (*Smooth Median Absolute Deviation*) proposé dans [13] qui correspond à une estimation robuste de la variance :

$$\text{SMAD}(\mathbf{f}_g, \mathbf{f}_d) = \sum_{k=0}^{h-1} \left(|\mathbf{f}_g - \mathbf{f}_d - \text{med}(\mathbf{f}_g - \mathbf{f}_d)|^2 \right)_{k:N_f-1}$$

où $k:N_f-1$ indique que les valeurs des carrés des écarts à la médiane sont ordonnées. On prend la somme des h plus petites. Ainsi, les plus grandes différences entre les deux ensembles ne sont pas prises en compte car on suppose que ce sont les pixels occultés qui les provoquent. SMAD mesure une dissimilarité alors que ZNCC mesure une similarité. Pour faciliter les comparaisons, nous décidons d'utiliser $-\text{SMAD}$.

Sélection. Les candidats qui ont donné le meilleur score de corrélation doivent satisfaire plusieurs contraintes afin d'être sélectionnés. Tout d'abord une contrainte de région est utilisée. Ainsi, pour le germe considéré, nous ne propageons que ses voisins appartenant à la même région. En effet, si le voisin d'un germe n'appartient pas à la même région de profondeur que celui-ci, alors, la probabilité qu'il ait une valeur de disparité très différente de

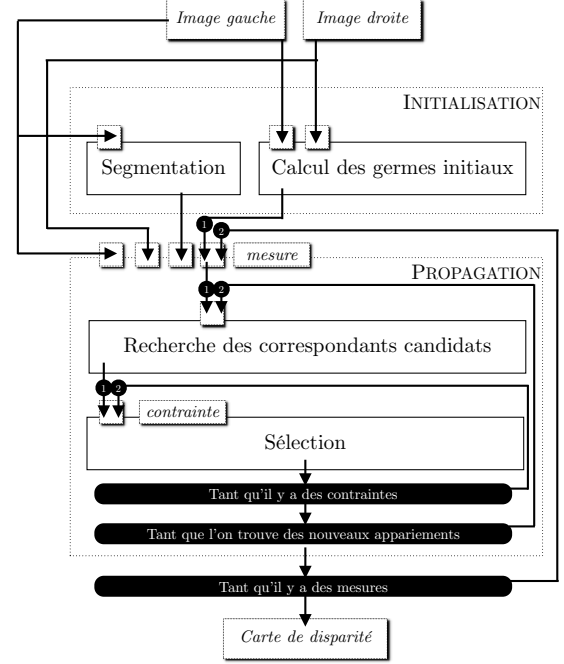


FIG. 2 – Ce schéma résume les principales étapes de la méthode proposée. Les opérations sont marquées dans des rectangles, les paramètres sont notés en italique dans les boîtes ombrées et les conditions sont dans les boîtes noires arrondies. Pour nos tests nous avons utilisé plusieurs segmentations, un ensemble théorique de germes, deux mesures (ZNCC puis SMAD) et deux contraintes (contrainte de seuil et vérification bidirectionnelle).

celle du germe est plus forte. Dans ce cas, son correspondant n'est donc pas au voisinage du correspondant du germe. Ensuite, une contrainte de seuil est appliquée. Les couples dont le score de corrélation est trop faible sont rejetés. Enfin, nous décidons pour chaque appariement candidat d'effectuer une vérification bidirectionnelle au moment de la validation (voir l'algorithme 2.1 point 2.2.2). Si la contrainte n'est pas vérifiée, nous ne validons pas l'appariement. Le pixel considéré reste donc sans correspondant et peut donc être traité à nouveau avec d'autres candidats (induits par d'autres germes) lors de l'itération suivante.

4 Protocole d'évaluation

Dans un premier temps, nous réalisons des tests pour savoir comment choisir nos seuils de façon judicieuse pour la méthode en deux phases. Dans un second temps, nous comparons les résultats des différentes méthodes.

4.1 Critères

Afin de prendre en compte les zones de discontinuités et d'occultations, nous reprenons les critères proposés dans [2]. Soit d_{ref} la fonction de disparité théorique et occ la disparité symbolique attribuée à un pixel occulté, nous posons :

$$\text{Err}_{i,j}^g = \|d(\mathbf{p}_{i,j}^g) - d_{\text{ref}}(\mathbf{p}_{i,j}^g)\|. \quad (1)$$

Le calcul de $\text{Err}_{i,j}^g$ n'a de signification que si $d(\mathbf{p}_{i,j}^g) \neq \text{occ}$ et $d_{\text{ref}}(\mathbf{p}_{i,j}) \neq \text{occ}$. Les critères étudiés sont les pourcen-

tages suivants :

- CORR : appariements corrects, un appariement est correct si $\text{Err}_{i,j}^g < 2$.
- ERREUR : appariements erronés, un appariement est erroné si $\text{Err}_{i,j}^g \geq 2$.

La densité (DENS) est le pourcentage de pixels pour lesquels une méthode a donné une réponse (pixel apparié ou occulté).

4.2 Images testées

Les couples que nous utilisons pour nos tests font partie des couples fournis par [6, 14, 15, 16]. Pour la plupart, les vérités terrain sont obtenues en utilisant une technique de projection de lumière structurée. Nous décidons de distinguer quatre catégories de couples présentant chacune un problème spécifique. Pour chaque catégorie, nous sélectionnons trois couples d'images significatifs :

- CAT. 1 : images avec de larges zones homogènes (*Bowling1*, *Plastic*, *Midd2*);
- CAT. 2 : images avec beaucoup de zones occultées et de nombreuses discontinuités (*Cones*, *Dolls*, *Moebius*);
- CAT. 3 : images avec des textures répétitives (*Aloe*, *Cloth1*, *Cloth4*);
- CAT. 4 : images mixtes (*Art*, *Venus*, *Teddy*).

Afin de ne juger que de la qualité des différentes approches de propagation, nous partons dans tous les cas d'un même ensemble initial de germes qui ont été mis en correspondance en utilisant les disparités théoriques.

4.3 Évaluations

Nous réalisons deux évaluations :

- ÉVAL. 1 porte sur les pixels ayant donné une réponse ;
- ÉVAL. 2 porte sur la totalité de l'image.

Ces deux évaluations sont complémentaires pour quantifier d'une part le gain de fiabilité de la méthode et d'autre part l'impact sur les performances globales, voir 5.3.

4.4 Influence de la segmentation initiale

Nous souhaitons évaluer l'apport de la segmentation initiale pour la méthode proposée. Pour cela, nous partons de plusieurs segmentations. Le programme EDISON² est utilisé. Nous faisons varier les paramètres afin d'obtenir les quatre segmentations suivantes :

- 0 : pas de segmentation ;
- SEG. 1 : sous-segmentation (une région peut contenir des pixels de profondeurs différentes) ;
- SEG. 2 : segmentation moyenne (une région contient des pixels de même profondeur) ;
- SEG. 3 : sur-segmentation (plusieurs petites régions connexes ont la même profondeur).

L'appréciation des différentes segmentations est visuelle.

4.5 Influence des seuils

Comment devons-nous choisir nos seuils sur les scores de corrélation obtenus lors des deux phases ? Un seuil strict

sur la mesure ZNCC va limiter la propagation de la première phase laissant ainsi plus de disparités à calculer lors de la seconde phase plus spécifique aux zones proches des discontinuités. Nous devons trouver ici un compromis entre les deux mesures. Nous faisons donc varier différents seuils et réalisons nos mesures pour juger du résultat. Pour ZNCC, nous fixons trois seuils :

- SEUIL PERMISSIF : 0.5 ;
- SEUIL INTERMÉDIAIRE : 0.75 ;
- SEUIL STRICT : 0.95.

Pour SMAD, nous fixons deux seuils :

- SEUIL PERMISSIF : -200 ;
- SEUIL STRICT : -50.

La mesure SMAD n'étant pas normalisée (contrairement à la mesure ZNCC), ces seuils sont déterminés après avoir observé que la majorité des couples validés par une mesure SMAD donnent une valeur située autour de -100.

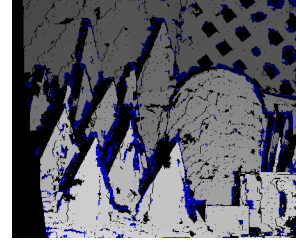


FIG. 3 – Les disparités en gris correspondent à celles calculées lors de la première phase avec ZNCC. Sur cet exemple, elles représentent 71% de l'image. Celles en bleu correspondent aux disparités calculées lors de la seconde phase avec SMAD. Sur cet exemple, elles représentent 13% de l'image. Les parties noires sont celles qui ne sont pas calculées.

4.6 Comparaison des différentes approches

Nous souhaitons comparer et évaluer les différentes approches dont nous parlons dans cet article. Pour l'approche séquentielle, nous implémentons la méthode de propagation de LHUILLIER présentée dans [9] et celle de MA présentée dans [11]. Pour l'approche simultanée nous implémentons la méthode de propagation de CHEN présentée dans [3]. Nous les comparons également avec la méthode MULTI-MESURE proposée dans cet article. Toujours à titre de comparaison, nous évaluons nos critères en faisant une mise en correspondance locale classique avec les mesures de corrélation ZNCC, d'une part, et SMAD, d'autre part.

5 Résultats et analyses

5.1 Influence de la segmentation initiale

Les résultats résumés par la figure 4 montrent que l'utilisation d'une segmentation permet d'améliorer les résultats pour les CAT. 1 & 4. On remarque également une légère augmentation des performances pour la CAT. 2. En revanche, on ne constate pas d'amélioration pour la CAT. 3. Cela illustre bien l'utilité d'une segmentation lorsque l'image contient des zones homogènes limitant ainsi les erreurs de propagation. Dans tous les cas, on constate que l'utilisation d'une sur-segmentation dégrade

²www.caip.rutgers.edu/riul/research/code/EDISON/

les performances. Pour mieux comprendre ce phénomène, nous pensons qu’une étude plus approfondie de la relation entre les germes initiaux et les régions est nécessaire mais par manque de place cette étude ne fait pas l’objet de cette publication.

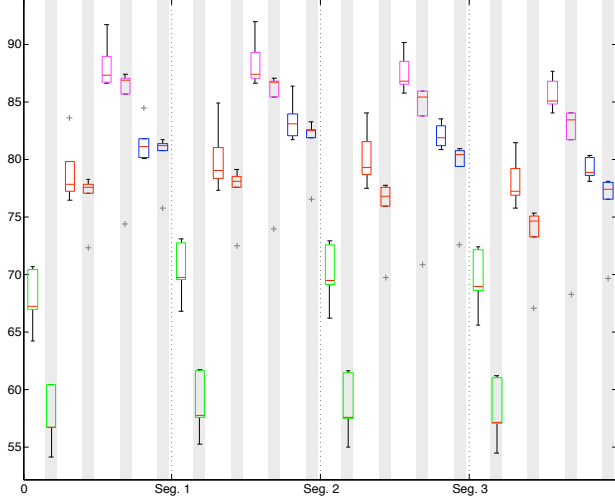


FIG. 4 – Cette figure montre les différents pourcentages d’appariements corrects (COR) obtenus en utilisant différentes segmentations initiales (0, SEG. 1, 2 & 3). Les quatre catégories sont présentées dans l’ordre et distinguées par les couleurs suivantes : CAT. 1 : vert, CAT. 2 : rouge, CAT. 3 : magenta et CAT. 4 : bleu. Pour chacune, il y a deux résultats. Le premier correspond au résultat de ÉVAL. 1. Le second, dans les colonnes grisées, correspond aux résultats de ÉVAL. 2. Pour chaque catégorie et chaque évaluation, on fait varier les seuils sur les scores de corrélation (voir 4.1). Ainsi, nous avons six résultats différents pour chaque combinaison. Ces résultats sont présentés sous la forme de diagrammes en forme de boîtes. Cette représentation permet d’avoir une vue d’ensemble de la distribution de tous les résultats. Elle montre les valeurs minimales et maximales (traits noirs), la valeur médiane (trait rouge) et les premier et troisième quartiles (début et fin des boîtes). Certaines données trop éloignées des autres sont considérées comme atypiques et sont marquées d’une croix grise.

5.2 Influence des seuils

Les résultats de nos tests pour le choix des seuils sur les mesures employées par la méthode en deux phases sont présentés pour l’ensemble des couples testés dans le tableau 1 qui distingue les quatre catégories d’images.

De façon générale, on constate qu’il est préférable d’avoir des seuils plus stricts si on cherche à obtenir un maximum d’appariements corrects. En revanche, si le résultat recherché est un compromis entre COR et DENS, pour les CAT. 2, 3 & 4, on constate qu’il semble plus intéressant d’utiliser un seuil strict sur ZNCC (0.95) et plus tolérant sur SMAD (-200) qu’un seuil permissif sur ZNCC (0.5) et strict sur SMAD (-50, voir tableau 1). Cela justifie bien l’intérêt de l’utilisation d’une mesure robuste en deuxième passe.

5.3 Comparaison des différentes approches

Dans cette section nous tentons de faire une comparaison des différentes méthodes présentées en essayant de prendre en compte le mieux possible les spécificités de

SEUILS	CAT. 1	CAT. 2	CAT. 3	CAT. 4
0.95/-50	67.33 (62.52)	84.92 (70.82)	92.05 (65.36)	86.43 (74.75)
0.95/-200	66.85 (67.81)	79.71 (87.05)	87.53 (86.97)	83.50 (89.01)
0.75/-50	69.72 (66.15)	81.09 (82.15)	89.30 (80.92)	84.01 (84.37)
0.75/-50	69.58 (69.10)	78.50 (88.17)	87.08 (87.31)	82.11 (89.83)
0.5/-50	73.08 (70.02)	78.40 (86.95)	87.38 (85.29)	82.82 (87.94)
0.5/-200	72.81 (72.46)	77.30 (89.49)	86.69 (87.98)	81.77 (90.87)

TAB. 1 – Ce tableau présente les pourcentages d’appariements corrects (COR) obtenus par la méthode MULTI-MESURE en faisant varier les seuils sur ZNCC/SMAD et en partant de la segmentation initiale SEG. 1. On marque entre parenthèses la densité DENS obtenue pour chaque résultat. Pour chaque catégorie, on note en gras les meilleurs résultats et les meilleures densités.

chacune. Le tableau 2 représente les résultats obtenus pour les différentes méthodes. On constate que la méthode de LHUILLIER donne les meilleurs résultats mais donne un résultat moins dense que les autres méthodes, en particulier pour la CAT. 1. Cela s’explique par la contrainte d’homogénéité qui est utilisée par cette méthode. En revanche, la méthode de MA, n’imposant quasiment pas de contraintes mise à part la contrainte de région, donne des résultats plus denses mais de moins bonne qualité. La méthode MULTI-MESURE donne un bon compromis entre densité et pourcentage d’appariements corrects.

MÉTHODE	CAT. 1	CAT. 2	CAT. 3	CAT. 4
CHEN	64.59 (59.22)	75.46 (85.16)	85.63 (84.21)	78.15 (82.68)
LHUILLIER	84.94 (22.49)	86.68 (66.64)	93.07 (78.98)	83.68 (58.40)
MA	64.26 (98.96)	71.31 (94.77)	83.11 (93.10)	76.35 (94.97)
ZNCC	54.71	78.77	88.78	79.67
SMAD	45.21	82.71	91.40	78.91

TAB. 2 – Ce tableau présente les pourcentages d’appariements corrects COR pour les autres méthodes testées : CHEN, LHUILLIER & MA. On marque également entre parenthèses la densité DENS obtenue pour chaque résultat. Toujours à titre de comparaison on note les résultats obtenus avec les méthodes de mise en correspondance locale par corrélation avec les mesures ZNCC et SMAD. Dans ce cas, le résultat est dense.

Nous remarquons également que la densité moyenne que nous obtenons avec la méthode MULTI-MESURE, la méthode de MA et la méthode de CHEN est de 84.28%. Les zones non calculées correspondent aux pixels ne vérifiant pas les contraintes utilisées. Par conséquent, dans ce cas, nous supposons que ces pixels pour lesquels nous n’avons pas trouvé de correspondant n’en ont pas et sont donc occultés. Nous décidons de recalculer les pourcentages d’appariements corrects en considérant tous les

pixels de l'image. Les résultats de cette nouvelle évaluation, ÉVAL. 2, sont présentés dans la figure 4 (colonnes grisées). Cette figure montre que notre hypothèse se justifie pour toutes les catégories sauf la CAT. 1. En effet, la perte d'appariements corrects ne semble pas être significative pour les CAT. 2, 3 & 4.

Cette hypothèse semble également être valable pour les méthodes de MA et CHEN. Les résultats moyens sont présentés dans le tableau 3.

MÉTHODE	ÉVAL. 1	ÉVAL. 2
MULTI-MESURE	79.64 (85.2)	77.34 (100)
CHEN	75.95 (77.81)	72.11 (100)
MA	73.75 (95.45)	74.16 (100)

TAB. 3 – ÉVAL. 1 correspond à l'évaluation sur les parties calculées de l'image. ÉVAL. 2 correspond à l'évaluation sur la globalité de l'image. Ce tableau présente les résultats moyens pour toutes les catégories. Les résultats pris en compte pour la méthode MULTI-MESURE sont obtenus pour ÉVAL. 2 avec les seuils qui donnent les meilleurs résultats moyens (0.5/-200). Il s'agit des seuils les plus permissifs ayant donné la meilleure densité lors de la propagation.

6 Conclusion et perspectives

Les contributions de ce papier portent à trois niveaux différents. Premièrement, nous avons donné un état de l'art des différentes approches de mise en correspondance par propagation de germes dans le voisinage direct. Deuxièmement, nous avons apporté une amélioration à l'approche simultanée. Cette amélioration se déroule en deux phases. Une phase faisant intervenir une mesure classique suivie d'une phase faisant appel à une mesure robuste dans les zones proches des discontinuités. Afin d'améliorer encore cette robustesse, la propagation est guidée par une segmentation couleur. Nous avons également utilisé une contrainte de vérification bidirectionnelle au fur et à mesure de la propagation. Troisièmement, nous proposons une étude comparative des méthodes proposées avec les méthodes existantes. Les résultats obtenus ont montré que l'amélioration que nous proposons permet d'obtenir un bon compromis entre la qualité du résultat et sa densité.

Cependant, pour cette méthode, la sélection des germes doit avoir une influence sur le résultat final. C'est la raison pour laquelle notre prochaine étude portera sur cet aspect. De plus, il nous semble possible et intéressant d'étendre cet algorithme en utilisant des ensembles de germes initiaux différents et en fusionnant ensuite les résultats obtenus. Cette technique pourrait être utilisée notamment pour étendre notre algorithme aux images en couleur.

Références

[1] M. Z. Brown, D. Burschka et G. D. Hager. « Advances in Computational Stereo ». *PAMI*, 25(8):993–1008, 2003.

[2] S. Chambon et A. Crouzil. « Dense matching using correlation: new measures that are robust near occlusions ». Dans *BMVC*, volume 1, pages 143–152, 2003.

[3] Q. Chen et G. Medioni. « A Volumetric Stereo Matching Method: Application to Image-Based Modeling ». Dans *CVPR*, volume 1, pages 1029–1034, 1999.

[4] D. Comaniciu et P. Meer. « Robust analysis of feature spaces: color image segmentation ». Dans *CVPR*, pages 750–755, 1997.

[5] C. Harris et M. Stephens. « A Combined Corner and Edge Detector ». Dans *Alvey Vision Conference*, pages 147–151, 1988.

[6] H. Hirschmuller et D. Scharstein. « Evaluation of Cost Functions for Stereo Matching ». Dans *CVPR*, pages 1–8, 2007.

[7] J. Kannala et S. S. Brandt. « Quase-Dense Wide Baseline Matching Using Match Propagation ». Dans *CVPR*, 2007.

[8] A. Klaus, M. Sormann et K. Karner. « Segment-Based Stereo Matching Using Belief Propagation and a Self-Adapting Dissimilarity Measure ». Dans *ICPR*, volume 3, pages 15–18, 2006.

[9] M. Lhuillier et L. Quan. « Match Propagation for Image-Based Modeling and Rendering ». *PAMI*, 24(8):1140–1146, 2002.

[10] D. G. Lowe. « Object Recognition from Local Scale-Invariant Features ». Dans *ICCV*, volume 2, pages 1150–1157, 1999.

[11] Y. Z. Ma et W. J. Liu. « Progressive Matching Based on Segmentation for 3D reconstruction ». Dans *International Conference on Computer and Information Technology*, pages 575–579, 2005.

[12] K. Mikolajczyk et C. Schmid. « Scale & Affine Invariant Interest Point Detectors ». *IJCV*, 60(1):63–86, 2004.

[13] P. J. Rousseeuw et C. Croux. *L₁-Statistical Analysis and Related Methods*. Dans Y. Dodge, éditeur, *Explicit Scale Estimators with High Breakdown Point*, pages 77–92. Elsevier, 1992.

[14] D. Scharstein et C. Pal. « Learning conditional random fields for stereo ». Dans *CVPR*, 2007.

[15] D. Scharstein et R. Szeliski. « A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms ». *IJCV*, 47(1):7–42, 2002.

[16] D. Scharstein et R. Szeliski. « High-Accuracy Stereo Depth Maps Using Structured Light ». Dans *CVPR*, volume 1, pages 195–202, 2003.

[17] R. Szeliski et D. Scharstein. « Symmetric Sub-Pixel Stereo Matching ». Dans *ECCV*, volume 2, pages 525–540, 2002.

[18] Y. Wei et L. Quan. « Region-based Progressive Stereo Matching ». Dans *CVPR*, volume 1, pages 106–113, 2004.

[19] H. Zhang et S. Negahdaripour. « Fast and robust progressive stereo reconstruction by symmetry guided fusion ». *Oceans – Europe*, 1:551–556, 2005.

[20] Z. Zhang et Y. Shan. « A Progressive Scheme for Stereo Matching ». Dans *European Workshop on 3D Structure from Multiple Images of Large-Scale Environments*, volume 2018 de *LNCS*, pages 68–85, 2000.